

GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 15:47:28 ; Search time 37.5 Seconds (without alignments)

54,710 Million cell updates/sec

Title: US-09-331-631A-22-COPY\_25\_84

Sequence: 1 EDDNNHHHGKSGCVRRC..... EKQERSRHEADRSSEGSS 60

Perfect score: 350

Scoring table: BL2SUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SIDS1/gcdata/geneseq/geneseq/AA1980.DAT:\*

2: /SIDS1/gcdata/geneseq/geneseq/AA1981.DAT:\*

3: /SIDS1/gcdata/geneseq/geneseq/AA1982.DAT:\*

4: /SIDS1/gcdata/geneseq/geneseq/AA1983.DAT:\*

5: /SIDS1/gcdata/geneseq/geneseq/AA1984.DAT:\*

6: /SIDS1/gcdata/geneseq/geneseq/AA1985.DAT:\*

7: /SIDS1/gcdata/geneseq/geneseq/AA1986.DAT:\*

8: /SIDS1/gcdata/geneseq/geneseq/AA1987.DAT:\*

9: /SIDS1/gcdata/geneseq/geneseq/AA1988.DAT:\*

10: /SIDS1/gcdata/geneseq/geneseq/AA1989.DAT:\*

11: /SIDS1/gcdata/geneseq/geneseq/AA1990.DAT:\*

12: /SIDS1/gcdata/geneseq/geneseq/AA1991.DAT:\*

13: /SIDS1/gcdata/geneseq/geneseq/AA1992.DAT:\*

14: /SIDS1/gcdata/geneseq/geneseq/AA1993.DAT:\*

15: /SIDS1/gcdata/geneseq/geneseq/AA1994.DAT:\*

16: /SIDS1/gcdata/geneseq/geneseq/AA1995.DAT:\*

17: /SIDS1/gcdata/geneseq/geneseq/AA1996.DAT:\*

18: /SIDS1/gcdata/geneseq/geneseq/AA1997.DAT:\*

19: /SIDS1/gcdata/geneseq/geneseq/AA1998.DAT:\*

20: /SIDS1/gcdata/geneseq/geneseq/AA1999.DAT:\*

21: /SIDS1/gcdata/geneseq/geneseq/AA2000.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	350	100.0	593	19	W62835	RESULT 1
2	129.5	37.0	637	19	W62835	W62835 standard; Protein; 593 AA.
3	80.5	23.0	625	19	W62830	AC W62835;
4	79.5	22.7	666	19	W62829	XX DT 27-OCT-1998 (first entry)
5	77.5	22.1	666	19	W62828	XX DE Zea mays antimicrobial protein.
6	70	20.0	2337	19	W37878	XX KW antimicrobial protein; infestation; control.
7	68.5	19.6	185	20	Y60129	XX OS Zea mays.
8	68.5	19.6	2251	16	R71009	XX PN W09827805-A1.
9	68.5	19.6	2270	16	R71010	XX PD 02-JUL-1998.
10	67.5	19.3	132	20	Y07004	XX PF 22-DEC-1997; 97WO-AU00874.
11	67.5	19.3	432	20	W93954	XX PR 20-DEC-1996; 96AU-0004275.
12	67	19.1	525	19	W62831	XX PA (BETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
					PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;	
					XX DR WPI; 1998-377279/32.	
					XX PT Novel anti-microbial protein from e.g. Macadamia integrifolia - useful for controlling microbial infestations of plants or mammals	
					XX PS Claim 1; Page 58-60; 96pp; English.	
					XX CC The sequence is that of an antimicrobial protein which can be used to control microbial infestations in plants and mammalian	
					CC The sequence is that of an antimicrobial protein which can be used to control microbial infestations in plants and mammalian	

Sequence encoded b Gossypium hirsutum Sequence of the al Human podocalyxin- Human calcium chan Human neuronal chan Human calcium chan Calcium channel al Sequence of the al Human SH3D1A prote Human SH3D1A prote Calcium channel al HIV C76S protein s Murine PGC-1 prote Human SC46 protein Human neuronal cal Human transmembran Membrane-bound pro Amino acid sequenc Sequence of clone Sequence encoded b Biorythm marker p G. max truncated s G. max SBP1 protei Renal cancer assoc HIV E21.24P protei HIV E21.24P protei HIV L68S protein s

Ze a mays. Human endometrium Human neuronal cal Human neuronal cal Breast cancer asso Human regulatory m Theobroma cacao an



XX	22-DEC-1997;	97WO-AU00874.
PR	20-DEC-1996;	96AU-0004275.
XX	PA (RETR-)	COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX	PT	Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
XX	DR	WPI; 1998-377279/32.
XX	N-PSDB;	V42311.
PT	Novel anti-microbial protein from e.g. Macadamia integrifolia - useful for controlling microbial infestations of plants or mammals	
XX	PT	Novel anti-microbial protein from e.g. Macadamia integrifolia - useful for controlling microbial infestations of plants or mammals
PS	Claim 1; Page 34-36; 96pp; English.	
XX	CC	The sequence is that of an antimicrobial protein which can be used to control microbial infestations in plants and mammalian animals.
CC	CC	be used to control microbial infestations in plants and mammalian animals.
CC	CC	be used to control microbial infestations in plants and mammalian animals.
XX	SQ	Sequence 666 AA;
Query Match	22.7%	Score 79.5; DB 19; Length 666;
Best Local Similarity	32.8%	Pred. No. 0.18; Pred. No. 0.18;
Matches	20;	Conservative 10; Mismatches 26; Indels 5; Gaps 3;
Qy	1	EDDNHHHHGGHKGSGQCVRR--EDRPHWQRRCLEOC-REEREKQERSRHEADDNSG 56
CC	114	eeynqrdrpqgqyeqcaercqrheteprhmq tcqqrcerrvekektrkqkryeqqred 172
Db	57	E 57
Qy	173	e 173
PT	RESULT 5	
XX	W62828	W62828 standard; Protein; 666 AA.
ID	W62828	
XX	AC	W62828;
AC	AC	W62828;
XX	DT	27-OCT-1998 (first entry)
DE	DE	Macadamia integrifolia antimicrobial protein.
XX	DE	antimicrobial protein; infestation; control.
OS	OS	Macadamia integrifolia.
FH	Key	Location/Qualifiers
FT	Peptide	1..28
FT	/note="signal peptide"	
FT	Protein	29..666
FT	/note="mature protein"	
XX	PN	W09827805-A1.
XX	PD	02-JUL-1998.
XX	PF	22-DEC-1997; 97WO-AU00874.
XX	PR	20-DEC-1996; 96AU-0004275.
XX	PA	(RETR- ) COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX	PT	Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
XX	DR	WPI; 1998-377279/32.
XX	N-PSDB;	V42310.
PT	Novel anti-microbial protein from e.g. Macadamia integrifolia - useful for controlling microbial infestations of plants or mammals	
XX	PT	Novel anti-microbial protein from e.g. Macadamia integrifolia - useful for controlling microbial infestations of plants or mammals
PS	Claim 1; Page 34-36; 96pp; English.	
XX	CC	The sequence is that of an antimicrobial protein which can be used to control microbial infestations in plants and mammalian animals.
CC	CC	be used to control microbial infestations in plants and mammalian animals.
CC	CC	be used to control microbial infestations in plants and mammalian animals.
XX	SQ	Sequence 666 AA;
Query Match	22.1%	Score 77.5; DB 19; Length 666;
Best Local Similarity	31.1%	Pred. No. 0.31; Pred. No. 0.31;
Matches	19;	Conservative 12; Mismatches 25; Indels 5; Gaps 3;
Qy	1	EDDNHHHHGGHKGSGQCVRRCEDR--PHWQRRCLEOC-REEREKQERSRHEADDNSG 56
CC	114	eeynqrdrpqgqyeqcaercqrheteprhmq tcqqrcerrvekektrkqkryeqqred 172
Db	57	E 57
Qy	173	e 173
PT	RESULT 6	
XX	W37878	W37878 standard; Protein; 2337 AA.
ID	W37878	
XX	AC	W37878;
XX	DT	28-AUG-1998 (first entry)
DE	DE	Human calcium channel alB subunit.
KW	KW	Calcium channel; human; central nervous system disorder;
KW	KW	Lambert-Eaton syndrome; diagnosis; therapy.
XX	OS	Homo sapiens.
XX	XX	
PN	W09811131-A2.	
XX	PD	19-MAR-1998.
XX	PP	11-SEP-1997; 97WO-US16146.
XX	PR	16-SEP-1996; 96US-0713118.
XX	PA	(AMHP ) AMERICAN HOME PROD CORP.
XX	PI	Chen ARS, Franco R, Shuey DJ;
XX	XX	
DR	DR	WPI; 1998-207325/18.
XX	N-PSDB;	V29059.
PT	DNA encoding human neuronal calcium channel subunit(s) - useful for diagnosis of and treatment of central nervous system disorders, e.g. Lambert-Eaton syndrome	
PT	PT	This polypeptide comprises the alB subunit of the human neuronal calcium channel. cDNA clones (see v29059-61) encoding the alB subunit, the a2d subunit (see W37879) and a b3 subunit (see W37880) have been isolated. These have been inserted into expression vectors and are stably expressed in transformed cell lines. The transformed cells show omega-conotoxin GVIA toxin binding activity, and omega-conotoxin GVIA toxin sensitive potassium-stimulated calcium uptake, indicating that the proteins expressed by the clones are capable of forming a functioning calcium channel.
CC	CC	Nucleic acids encoding the 3 subunits, as well as vectors, host cells and methods of isolating nucleic acids encoding related calcium channels are disclosed. Fusion proteins incorporating related subunit proteins, antibodies, and assays for identifying agents
CC	CC	useful for controlling microbial infestations of plants or mammals

that modulate calcium channel activity are also provided. Such agents can be used to treat certain central nervous system disorders by altering calcium channel activity. Methods of diagnosing diseases associated with particular calcium channels, such as Lambert-Eaton syndrome, are disclosed.

```

Query Match          20.0%;  Score 70;  DB 19;  Length 2337;
Best Local Similarity 29.4%;  Pred. No. 8.6;  Mismatches 13;  Indels 40;  Gaps
Matches 25;  Conservative 7;  Mismatches 13;  Indels 40;  Gaps
Qy   4  NHHHHGHHKSGQVRCEDRPNHQPRCLEQ-----C 35
      :|||| :|| ||| :|| |:-----C 35
Db   2048 hhhhh----rchr-rdr---kqsllekpslsadmdgapssavvpglppgeqptqc 2096
Qy   36 REBERERKQOERSRHEADDRSGBCSS 60
      | :||||||| : :||| :-----C 35
Db   2097 r-rererrqrsrsqrrqpssss 2120

```

PA (META-) METAGEN GES GENOMFORSCHUNG MBH.  
 XX  
 PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;  
 XX  
 DR WPI: 1999-591957/51.  
 DR N-FSDB; Z42043.

XX  
 PT New nucleic acid sequences expressed in uterine cancer tissues, and  
 PT derived polypeptides, for treatment of uterine and endometrial cancer  
 PT and identification of therapeutic agents -  
 XX  
 PS Claim 23; Page 350-351; 444pp; German.

XX  
 CC This invention describes novel human nucleic acid (cDNA) sequences (A),  
 CC that are highly expressed in uterine tumour tissue and which have  
 CC anticancer and cytostatic activity. (A) are used (i) for recombinant  
 CC expression of polypeptides (B) and (ii) to isolate complete genes. (B)  
 CC are used (i) to identify agents suitable for treatment of uterine or  
 CC endometrial cancer; (ii) directly for treating these forms of cancer  
 CC (including expression from gene therapy vectors) and (iii) for generation  
 CC of specific antibodies. (A) are identified by assembling ESTs (expressed  
 CC sequence tags) from a particular tissue type before comparison of the  
 CC expression patterns. This allows a significantly longer fragment of the  
 CC gene to be revealed, so should reduce the number of failures associated  
 CC with the fact that ESTs from different libraries may represent different  
 CC parts of the same unknown gene, distorting the estimated frequency of  
 CC occurrence in a particular tissue. Y5941-1630328 represents protein derived  
 CC fragments encoded by the human endometrial tumour cDNA library derived

PS Claim 34; Page 201-211; 205pp; English.

XX  
 CC DNA encoding alpha 1E human calcium channel subunits have  
 CC been isolated from an oligo dT primed human hippocampus  
 CC library. The resulting clones, which are splice variants,  
 CC were designated alpha 1E-1 and alpha 1E-3. These splice variants  
 CC differ by virtue of a 57 bp insert in 1E-3. Alpha 1E-1 has  
 CC a calculated mol. wt. of 254,836 and alpha 1E-3 has a calculated  
 CC mol. wt. of 254,838. Alpha 1E-3 has a 19 AA insert relative to  
 CC alpha 1E-1 in the region that appears to be the cytoplasmic loop  
 CC between transmembrane domains IIS6 and IIS1.

XX  
 SQ Sequence 2251 AA:

Query	Match	Score	DB	Length
Best	Local Similarity	19.68;	16;	2251;
Matches	Conservative	31.4%;	Pred. No. 12;	
			Mismatches	29;
			Indels	13;
Oy	3	DNNHHHGKGSQCVRCEDRPHQRPRCLQCCRER-----EKRSRSR-----R		
Db	2006	dsghksdthpsgrerrrskerkhllspdvrcnseergtqadwesperrrspsegr		
Oy	50	#ADPSECFGS	59	



DT	30-JUN-1999	( first entry)
DE	Human regulatory molecule HRM-10 protein.	PN WO9827805-A1.
XX		XX
KW	Human regulatory molecule; HRM-10; cytostatic activity; immune modulator; transcription factor; enhancer; cell proliferation stimulation; cancer; treatment; microarray; detection; diagnosis; cell proliferation disease; leukemia; lymphoma; myeloma; adenocarcinoma; sarcoma; bladder; bone; brain; lung; liver; ovary; skin; teratocarcinoma; immune response; allergy; asthma; diabetes; multiple sclerosis; Grave's disease; myasthenia gravis.	XX
KW		XX
OS	Homo sapiens.	XX
XX	W09915658-A2.	XX
PN		XX
XX	01-APR-1999.	PD 02-JUL-1998.
PD		XX
XX	22-SEP-1998; 98WO-US19839.	PF 22-DEC-1997; 97WO-AU00874.
PF		XX
XX	22-SEP-1998; 98WO-US19839.	PR 20-DEC-1996; 96AU-0004275.
PR		XX
XX	23-SEP-1997; 97US-0933750.	PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PA		XX
XX	(INCY-) INCYTE PHARM INC.	PT Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP.
PA		XX
XX	PI Au-Young J, Bandman O, Guegler KJ, Hillman JL, Lai P, PT Shan P, Yue H;	DR WPI; 1998-377279/32.
XX	XX	XX
DR	WPI; 1999-254710/21.	PT Novel anti-microbial protein from e.g. Macadamia integrifolia - useful for controlling microbial infestations of plants or mammals.
DR	N-PSDB; X24068.	XX
XX		PS Claim 1; Page 47-49; 96pp; English.
PT	New human regulatory molecules	XX
XX		CC The sequence is that of an antimicrobial protein which can be used to control microbial infestations in plants and mammalian animals.
PS		XX
XX	PI Au-Young J, Bandman O, Guegler KJ, Hillman JL, Lai P, PT Shan P, Yue H;	SQ sequence 525 AA;
XX	XX	XX
CC	This invention describes novel human regulatory molecules (HRM) which have cytostatic activity and act as immune modulators, transcription factors or enhancers. The HRMs can be used to stimulate cell proliferation. Antagonists and agonists of the proteins of the invention can be used to treat cancer. The encoding nucleic acids can be used in microarrays to detect polynucleotides (and their expression levels) that encode HRMs in a biological sample. The HRMs and microarrays can be used to diagnose, treat or prevent cell proliferation diseases especially cancer, e.g. leukemia, lymphoma, myeloma, adenocarcinoma, sarcoma, cancer of e.g. bladder, bone, brain, lung, liver, ovary, skin, etc., teratocarcinoma, or multiple sclerosis, Grave's disease or myasthenia gravis.	Query Match Best Local Similarity 19.1%; Score 67; DB 19; Length 525; Matches 17; Conservative 5; Mismatches 12; Indels 4; Gap 4; Gap
CC		Qy 15 QCVRREDRPWHQPRPQQE-ER-EKRSRSR 50
CC		Db 44 qcqrceateeq--eqeqqrcreykeqqrqee 79
CC		RESULT 13
CC		R20181
CC		ID R20181 standard; Protein; 566 AA.
CC		XX
AC		AC R20181;
AC		XX
AC		DT 16-APR-1992 (first entry)
AC		XX
AC		DE Sequence encoded by 67 kD T. cacao protein cDNA.
AC		XX
AC		DE Cocoa; flavour; vicilin; seed storage protein.
AC		XX
AC		KW
AC		OS Theobroma cacao.
AC		XX
AC		PN WO919801-A.
AC		XX
AC		PD 26-DEC-1991.
AC		XX
AC		PR 07-JUN-1991; 91WO-GB00914.
AC		XX
AC		PR 11-JUN-1990; 90GB-0013016.
AC		XX
AC		PA (MRSC ) MARS UK LTD.
AC		XX
AC		PI Spencer ME, Hodge R, Deakin EA, Ashton S;
AC		XX
AC		DR WPI; 1992-024418/03.
AC		DR N-PSDB; Q20377.
AC		XX
AC		PT Recombinant cocoa proteins - are responsible for flavour in cocoa beans and produced in large quantities using yeast and bacterial expression vectors
AC		XX
AC		PT Claim 4; Fig 2; 59pp; English.
AC		XX
AC		CC The inventors claim a 67 kD and 31 kD T. cacao protein, and fragments, and encoding DNAs. The 47 kD and 31 kD proteins are
AC		CC
AC		OS Theobroma cacao.



